

Original Article

The Singer's Formant and Speaker's Ring Resonance: A Long-Term Average Spectrum Analysis

Sang-Hyuk Lee, MD · Hee-Jun Kwon, MD · Hyun-Jin Choi, MD · Nam-Hun Lee, MD
Sung-Jin Lee, MD · Sung-Min Jin, MD, PhD

Department of Otorhinolaryngology-Head & Neck Surgery, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Korea

Objectives. We previously showed that a trained tenor's voice has the conventional singer's formant at the region of 3 kHz and another energy peak at 8-9 kHz. Singers in other operatic voice ranges are assumed to have the same peak in their singing and speaking voice. However, to date, no specific measurement of this has been made.

Methods. Tenors, baritones, sopranos and mezzo sopranos were chosen to participate in this study of the singer's formant and the speaker's ring resonance. Untrained males (n=15) and females (n=15) were included in the control group. Each subject was asked to produce successive /a/ vowel sounds in their singing and speaking voice. For singing, the low pitch was produced in the chest register and the high notes in the head register. We collected the data on the long-term average spectra of the speaking and singing voices of the trained singers and the control groups.

Results. For the sounds produced from the head register, a significant energy concentration was seen in both 2.2-3.4 kHz and 7.5-8.4 kHz regions (except for the voices of the mezzo sopranos) in the trained singer group when compared to the control groups. Also, the chest register had a significant energy concentration in the 4 trained singer groups at the 2.2-3.1 kHz and 7.8-8.4 kHz. For speaking sound, all trained singers had a significant energy concentration at 2.2-5.3 kHz and sopranos had another energy concentration at 9-10 kHz.

Conclusion. The results of this study suggest that opera singers have more energy concentration in the singer's formant/speaker's ring region, in both singing and speaking voices. Furthermore, another region of energy concentration was identified in opera singer's singing sound and in sopranos' speaking sound at 8-9 kHz. The authors believe that these energy concentrations may contribute to the rich voice of trained singers.

Key Words. Speech acoustics

INTRODUCTION

Bartholomew stated that a good operatic voice needs a concentration of energy around 3 kHz (1). He also mentioned that this concentration must be produced with a special resonator in the larynx or lower pharynx. Sundberg offered a physiological explanation for this requirement. He constructed long-term average spectra of normal speech and orchestral music with and without a solo singer's voice (2-4). He demonstrated the presence of a secondary peak of acoustic energy around the 2-3 kHz region for operatic singing. It is now known as the singer's formant. He

also found that the singer's formant is a formant cluster of the third, fourth and fifth formants and that a prominent spectrum envelops the peak appearing in the vicinity of 3 kHz in all vowel spectra sung by male singers and alto singers.

A similar phenomenon observed in trained speakers has led investigators to examine the speaker's ring. Oliveira Barrichelo et al. found that singers show more energy concentration in the singer's formant/speaker's ring region in both singing and spoken vowels (5). Furthermore, it was observed that the singers' spoken vowel energy in the speaker's ring area was significantly larger than that of untrained speakers.

In a previous study, Titze and Jin showed that for trained tenors, in addition to the conventional singer's formant in the region of 3 kHz, another energy peak in the 8-9 kHz region was present (6). This peak was described as the second resonance of the epilarynx tube. Singers in other operatic voice ranges who produce vocal ringing are assumed to have the same peak in the singing and speaking voices.

• Received August 27, 2007
Accepted after revision October 4, 2007

• Corresponding author : **Sung-Min Jin, MD, PhD**
Department of Otorhinolaryngology-Head & Neck Surgery, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, 108 Pyeong-dong, Jongno-gu, Seoul 110-746, Korea
Tel : +82-2-2001-2266, Fax : +82-2-2001-2273
E-mail : strobojin@hanmail.net

Thus, the authors sought to investigate the energy peak in the singing voice around the 8-9 kHz region and in the speaking voice around the 3 and 8-9 kHz regions in 4 voice ranges of trained opera singers.

MATERIALS AND METHODS

We chose 15 tenors, 14 baritones, 7 sopranos and 5 mezzo sopranos in order to investigate the singer's formant. After 1 yr, we chose another 10 tenors, 15 baritones, 15 sopranos and 10 mezzo sopranos in order to investigate the speaker's ring. All subjects were enrolled at a music college, department of vocal music and they could reliably produce the head and chest registers. In the singer's formant group, age ranged from 23.3 to 34 yr and the average period of vocal training in singers' formant group was 4.6 yr for tenors, 5.4 yr for baritones, 8.8 yr for sopranos and 18 yr for mezzo sopranos. In the speakers' ring group the age ranged from 21.6 to 31.2 yr and the training period was 7.9 yr in tenors,

6.3 yr in baritones, 9.4 yr in sopranos and 15.6 yr in mezzo sopranos.

All voice samples were obtained in a special recording studio-room. Each subject was asked to produce a successive /a/ vowel for at least 3 sec at a high pitch (tenors, G4; baritones, E4; sopranos, F5; and mezzo sopranos, C5) and a low pitch (tenors, C3; baritones, D3; sopranos, D4; and mezzo sopranos, A3) and they were also asked to produce a sustained vowel /a/ as a usual speaking sound. The high pitch was produced in the head register and the low note in the chest register, as judged by a professor of music and a graduate of the conservatory.

The singing and spoken vowels were initially recorded on a SONY DAT recorder TCD-D7 (Sony, Tokyo, Japan). The uni-directional dynamic microphone was used for data acquisition at the standard distance between voice and microphone of 17 cm. The data were then stored in a computer. The sampling rate was 20 kHz. We adjusted the input level to the green LED lights on the front panel of the CSL external module in order to avoid overloading. The sound samples were analyzed with the

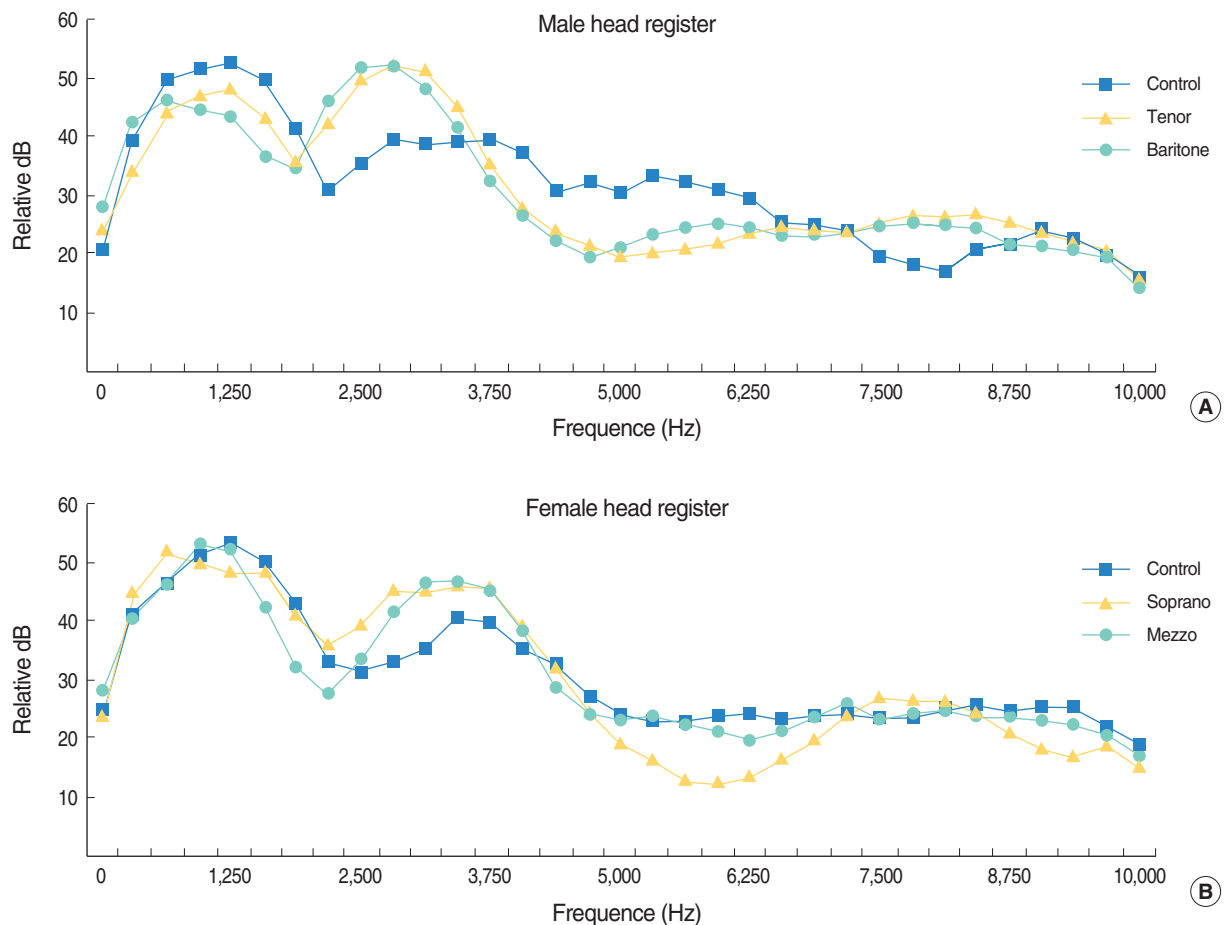


Fig. 1. Long term average spectrum of head register. (A) fifteen tenors, 14 baritones and male control group singing a prolonged /a/ vowel at a high pitch (tenor, G4; baritones, E4) in head voice. (tenor triangles, baritone circles, control squares). (B) seven sopranos, 5 mezzo sopranos and female control group singing a prolonged /a/ vowel at a high pitch (sopranos, F5; mezzo sopranos, C5) in head voice. (sopranos triangles, mezzo sopranos circles, control squares).

long-term average (LTA) power spectrum using the FFT algorithm of the Computerized Speech Lab (CSL, Model 4300B, Kay Elemetrics, Lincoln Park, NJ, USA). Long-term average frequency spectra were computed with a short analysis window (64 points or 3.2 ms). The energy level was recorded in dB at every 3.125 kHz.

The control group consisted of 15 male and 15 female adults who had no laryngeal abnormalities or abnormalities of articulation. Age ranged from 24.5 to 32.4 yr, which was similar with that of the trained singer's group. The control subjects were also asked to produce successive /a/ vowel sounds using their usual singing pitch (male, G4 and C3; female, F5 and D4) and usual speaking techniques.

Statistical analysis was performed using the Mann-Whitney test of the Statistical Package for Social Sciences (version 13.0, SPSS, Chicago, IL, USA). The results for each group were expressed as mean values \pm standard deviations. Statistical significance was defined as $P < 0.05$.

RESULTS

Data analysis included the evaluation of the difference in energy concentration along the frequency spectrum. Fig. 1 shows the energy levels of the singing vowel averages at a high pitch for each frequency analyzed for the tenor, baritone, soprano, mezzo soprano and control groups. A significant energy concentration was seen in the 2.2-3.4 kHz region ($P < 0.05$) and 7.5-8.4 kHz ($P < 0.05$) for the head registers of the tenors and baritones, respectively, compared to the untrained male control group (Fig. 1A). In addition, there were significant energy concentrations for the sopranos and mezzo sopranos compared to the untrained female control group in the 2.8-3.4 kHz region ($P < 0.05$). The sopranos showed significant energy concentration in the 7.8-8.4 kHz region ($P < 0.05$). However, the mezzo sopranos did not show energy concentration in the 8-9 kHz region (Fig. 1B).

Fig. 2 illustrates the energy levels of the singing vowel averages at a low pitch at each frequency analyzed for 4 categories of trained opera singers as well as the male and female control

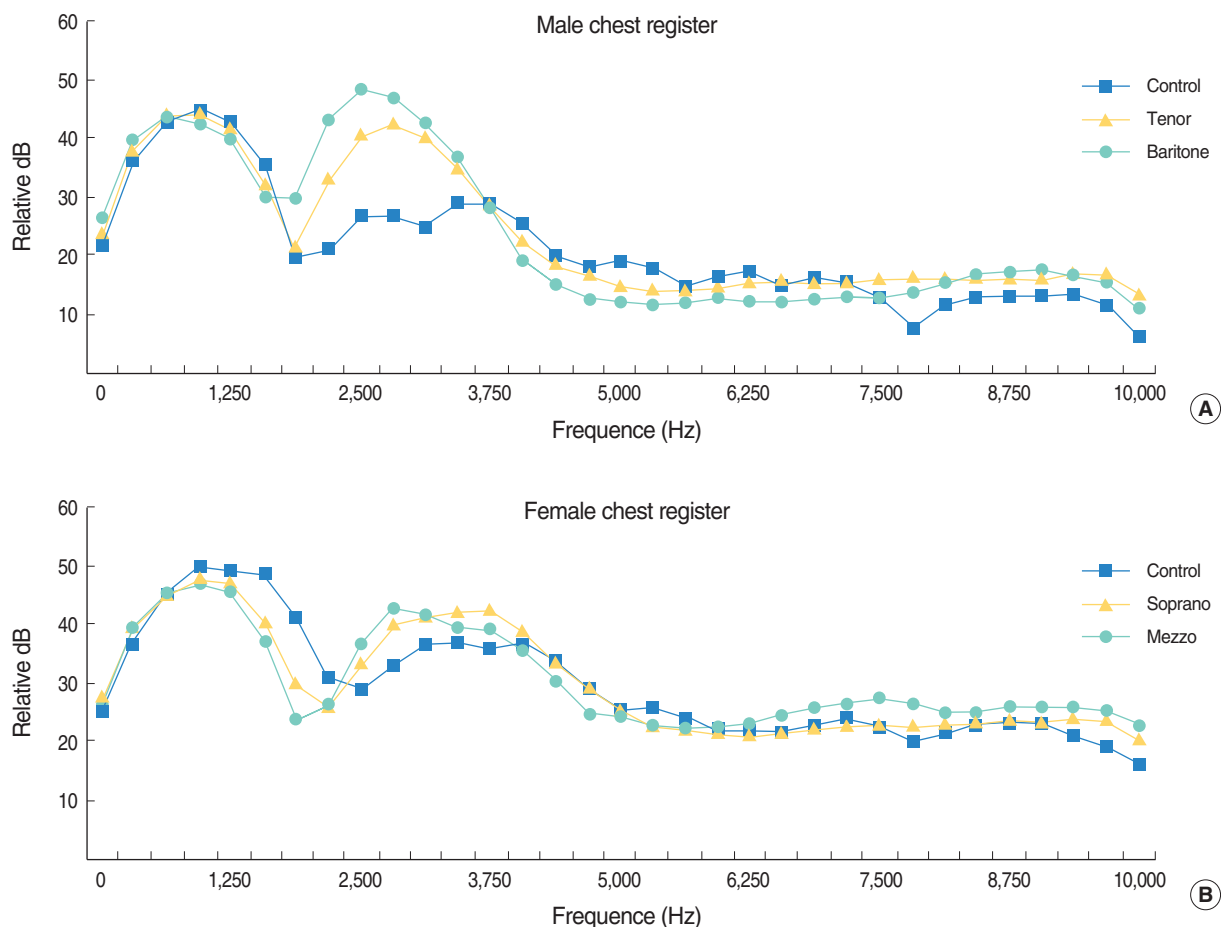


Fig. 2. Long term average spectrum of chest register. (A) fifteen tenors, 14 baritones and male control group singing a prolonged /a/ vowel at a low pitch (tenor, C3; baritones, D3) in chest voice. (tenor triangles, baritone circles, control squares). (B) seven sopranos, 5 mezzo sopranos and female control group singing a prolonged /a/ vowel at a low pitch (sopranos, D4; mezzo sopranos, A3) in chest voice. (sopranos triangles, mezzo sopranos circles, control squares).

groups. Similar to the head register sounds, a significant energy concentration was seen in the 2.2-3.1 kHz region ($P<0.05$) and the 7.8-8.4 kHz region ($P<0.05$) for the chest registers of the tenors and baritones compared to the untrained male control group (Fig. 2A). There were significant energy concentrations for the sopranos and mezzo sopranos compared to the untrained female control group in the 2.5-3.1 kHz region ($P<0.05$) and 7.8-8.4 kHz ($P<0.05$) (Fig. 2B).

The average energy levels of spoken vowel /a/ for each frequency which are analyzed for 4 categories of the trained opera singers and male/female control groups are shown in Fig. 3. A significant energy concentration was seen in the 2.2-3.4 kHz region ($P<0.05$) for the spoken vowel /a/ for the tenors and baritones compared to the untrained male control group (Fig. 3A). There was a significant energy concentration for the sopranos compared to the untrained female control group in the 2.5-5.3 kHz region ($P<0.05$) and the 9-10 kHz region ($P<0.05$). Mezzo sopranos showed significant energy concentration in the 2.2-3.7 kHz region ($P<0.05$) (Fig. 3B).

DISCUSSION

Voice characteristics are determined by the interactions of breathing mechanisms, the vibrating vocal folds and resonances of the vocal tract (7). The classically trained singer strives to develop a singer's formant that enhances the over all richness and ringing in his/her voice. The singer's formant is characterized by an increase in signal intensity between the third and fourth formants. This allows audiences to hear the singer's voice with amplification over the sounds of accompanying music (8). The greatest harmonics peak between 2 and 4 kHz is roughly the center frequency for the singer's formant (2). It makes the human voice easier to hear in the presence of a loud orchestral accompaniment (9). Many studies have described and analyzed the singer's formant (1-4, 7). Sundberg offered a physiological explanation, and he modeled the singer's formant with a small resonator inside the vocal tract in the epilaryngeal region (4).

Several researchers have made progress in the understanding of the speaker's formant and ring (5, 7, 10, 11). Recently, Oliveira Barrichelo et al. hypothesized that trained singers could carry

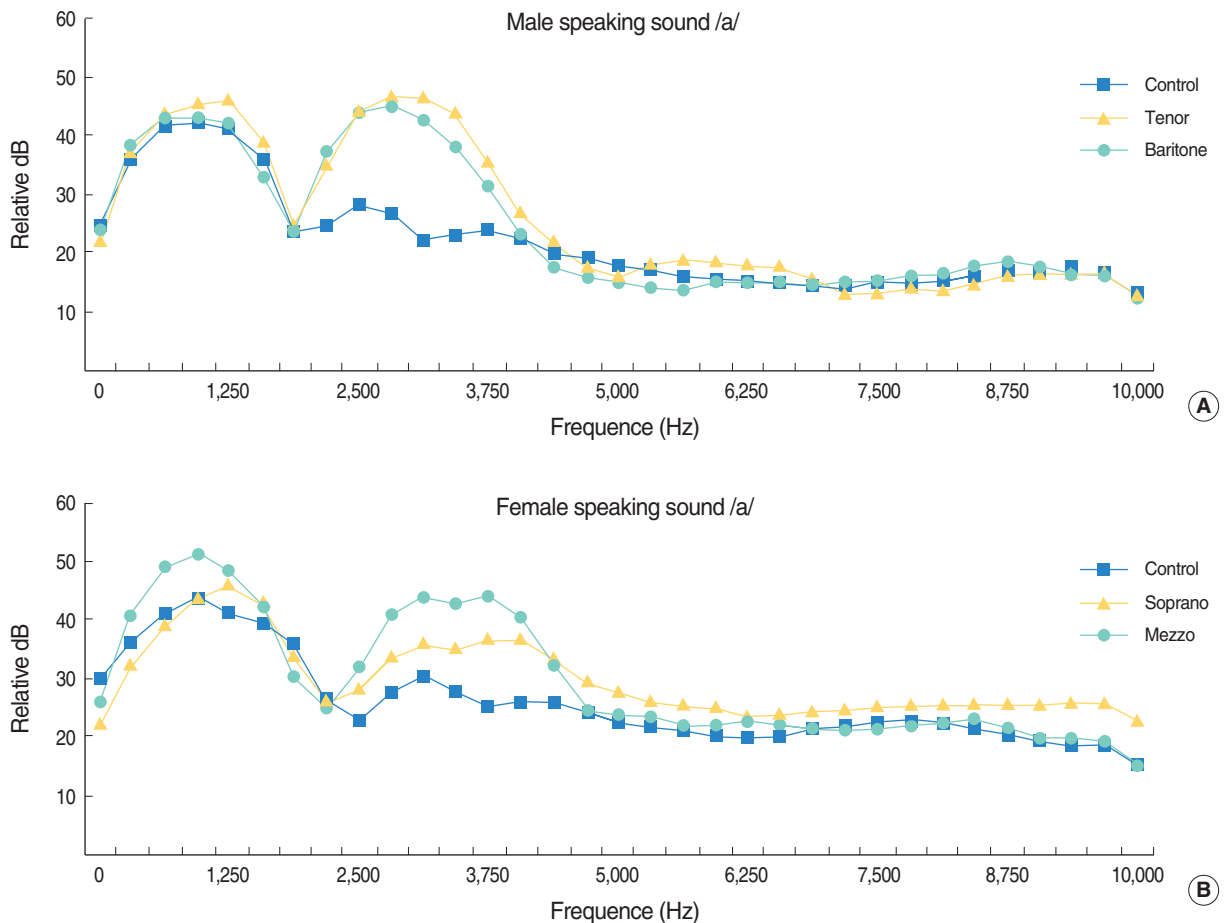


Fig. 3. Long term average spectrum of speaking sound. (A) ten tenors, 15 baritones and male control group speaking a prolonged /a/ vowel as an usual speaking sound. (tenor triangles, baritone circles, control squares). (B) fifteen sopranos, 10 mezzo sopranos and female control group speaking a prolonged /a/ vowel as an speaking sound. (sopranos triangles, mezzo sopranos circles, control squares).

their singing formant ability into normal speech (5). They found that singers' spoken vowel energy in the speaker's ring area was significantly larger than that of untrained speakers.

In this study, the authors investigated the energy peak in the singing voice around the 8-9 kHz region and in the speaking voice around the 3 kHz region in 4 categories of trained opera singers. We also found the energy peak in the 8-9 kHz region in sopranos' speaking sound.

Although we did not measure the sound pressure level during the recording of the singers' voice using a DAT recorder, we focused on the recording of the most outstanding voice as judged by a professor of music and a graduate of the conservatory. The data obtained extend and support previous findings in 2 areas. First, there was energy concentration around the 8 kHz region in most trained singers' singing voices. Second, the speaking voice of 4 categories of trained opera singers significantly demonstrated more energy around the 3 kHz region, and sopranos' speaking voice also showed another energy peak in the 8-9 kHz region.

We postulate that both the head/chest register and the speaking sounds, when formed in the classically trained singing technique, are closely related to the formation of the singer's formant/speaker's ring. In addition, it seemed that the singers had a tendency to produce the spoken vowel in a singing fashion as previously described by Stone et al. (7).

A significant finding of this study is that for trained operatic singers, in addition to the conventional singer's formant in the region of 3 kHz, another energy peak was observed in the region of 8-9 kHz. This second peak was interpreted as the second resonance of the epilarynx tube (6). We applied the formula that was derived for a uniform tube that was closed on one end (the glottis) and open on the other (the aryepiglottic collar). This formula is $F=(2n-1)c/4L$ where c is the speed of sound in the vocal tract (340 m/s), L is the length of the epilarynx tube (3 cm) and letter n is 2. Use of this formula revealed a second resonance frequency at about 8.5 kHz (6).

In this study, the authors found energy at around 8 kHz in most trained singers' singing voices and soprano' speaking voices. Although we did not measure the value in this study, there was the possibility that another energy peak could be theoretically present near the 15 kHz. However, the energy level around the 8-9 kHz region is about 20-30 dB below that of singer's formant region, suggesting that it may be beyond the range of perceived sound since the sensitivity of the human auditory system is also about 20 dB lower around the 8 kHz region instead of the 3 kHz region. However, with other recording and amplifi-

cation systems, where high frequency emphasis is given, a formant of 8-9 kHz could reach perceptual significance (6). It is thought that this second singers' formant may contribute to their rich and ringing voice.

Further understanding of this phenomenon will require additional studies involving some physical or physiological modeling. However, our findings suggest the possibility of a formerly unrecognized singer's formant in the 8-9 kHz region.

CONCLUSION

The results of this study suggest that singing voices of opera singers and some of their speaking voices may show more energy concentration in the singer's formant/speaker's ring region compared to the control group. Another energy concentration in the 8-9 kHz region was observed in all opera singers' singing vowel. We believe that this region is the second singer's/speaker's formant which can contribute to singers' rich voice.

REFERENCES

1. Bartholomew WT. A physical definition of 'good voice-quality' in the male voice. *J Acoust Soc Am*. 1934;6:25-33.
2. Sundberg J. Vocal tract resonance. In: Staloff RT, editor. *Professional Voice: The Science and Art of Clinical Care*. 2nd ed. San Diego: Singular Publishing Group, Inc; 1997. p.167-84.
3. Sundberg J. Articulatory interpretation of the "singing formant". *J Acoust Soc Am*. 1974 Apr;55(4):838-44.
4. Sundberg J. What's so special about singers? *J Voice*. 1990 Jun;4(2):107-19.
5. Oliveira Barrichelo VM, Heuer RJ, Dean CM, Sataloff RT. Comparison of singer's formant, speaker's ring, and LTA spectrum among classical singers and untrained normal speakers. *J Voice*. 2001 Sep;15(3):344-50.
6. Titze IR, Jin SM. Is there evidence of a second singer's formant? *J Singing*. 2003; 59(4):329-31.
7. Stone RE Jr, Cleveland TF, Sundberg J. Formant frequencies in country singers' speech and singing. *J Voice*. 1999 Jun;13(2):161-7.
8. Lundy DS, Roy S, Casiano RR, Xue JW, Evans J. Acoustic analysis of the singing and speaking voice in singing students. *J Voice*. 2000 Dec;14(4):490-3.
9. Sundberg J. Level and center frequency of the singer's formant. *J Voice*. 2001 Jun;15(2):176-86.
10. Nawka T, Anders LC, Cebulla M, Zurakowski D. The speaker's formant in male voices. *J Voice*. 1997 Dec;11(4):422-8.
11. Cleveland TF, Sundberg J, Stone RE. Long-term-average spectrum characteristics of country singers during speaking and singing. *J Voice*. 2001 Mar;15(1):54-60.